Small Nucleolar RNAs Tell a Different Tale

Kufel, Joanna; Grzechnik, Pawel

DOI:
10.1016/j.tig.2018.11.005

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Kufel, J & Grzechnik, P 2019, 'Small Nucleolar RNAs Tell a Different Tale', *Trends in Genetics*, vol. 35, no. 2, pp. 104-117. https://doi.org/10.1016/j.tig.2018.11.005

Trends in Genetics

CellPress
REVIEWS

Review

Small Nucleolar RNAs Tell a Different Tale

Joanna Kufel[1] and Pawel Grzechnik[2,*,@]

Transcribing RNA Polymerase II interacts with multiple factors that orchestrate maturation and stabilisation of messenger RNA. For the majority of noncoding RNAs, the polymerase complex employs entirely different strategies, which usually direct the nascent transcript to ribonucleolytic degradation. However, some noncoding RNA classes use endo- and exonucleases to achieve functionality. Here we review processing of small nucleolar RNAs that are transcribed by RNA Polymerase II as precursors, and whose 5′ and 3′ ends undergo processing to release mature, functional molecules. The maturation strategies of these noncoding RNAs in various organisms follow a similar pattern but employ different factors and are strictly correlated with genomic organisation of their genes.

Highlights

snoRNAs are functional, noncoding RNAs usually transcribed by RNA Polymerase II, the same enzyme that synthesises messenger RNAs. In contrast to mRNAs, snoRNAs are ribonucleolytically processed to achieve functionality.

snoRNA processing aims to erase mRNA hallmarks: the m^7G cap and the poly(A) tail. Different snoRNA maturation strategies are employed depending on organism and genomic organisation of snoRNA genes.

ncRNA-specific transcription termination factors recruit 3′ end processing exonucleases to independently transcribed pre-snoRNAs.

Removal of the m^7G cap is crucial for box C/D snoRNAs maturation and functionality in *S. cerevisiae*.

Alternative snoRNA processing may generate noncanonical functional snoRNAs.

Small Nucleolar RNAs: Essential and Functional Noncoding RNAs

During the past decade, it has become apparent that RNAs transcribed from noncoding regions of the genome play essential functions in various biological processes. This prompted researchers to gain an understanding of how noncoding RNA (ncRNA) biogenesis is controlled and whether (and how) this process resembles synthesis of messenger RNA (mRNA). Recent analyses of small nucleolar RNAs (snoRNAs) have provided fundamental information on how expression of these ncRNAs is regulated. SnoRNAs are essential, short (60–300 nt long), mostly nucleoli-localised, non-polyadenylated ncRNAs, present in all eukaryotic organisms. They are classified by the presence of highly conserved sequences ('boxes'), as either box C/D or box H/ACA (Figure 1A). The box C/D snoRNAs form a closed loop, which contains a box C and a box D (with conserved RUGAUGA and CUGA motifs, respectively), as well as a less conserved box C′ and box D′ [1]. The box H/ACA snoRNAs usually consist of two stem loops linked by the H box (ANANNA motif) and an ACA sequence near the 3′ end [1]. Box C/D snoRNAs with a long UG repeat and H/ACA snoRNAs with an additional CAB box (UGAG motif) are called small Cajal body-associated RNAs (scaRNA) (Figure 1A) and are localised in subnuclear structures known as Cajal bodies [2,3]. Both snoRNA classes are bound by a distinct set of proteins to form stable and catalytically active box C/D and box H/ACA ribonucleoprotein (snoRNP) structures [4].

The majority of snoRNAs carry specific sequences, which are complementary to other cellular RNAs and thus guide snoRNP to specific RNA substrates. The main function of canonical snoRNAs and scaRNAs is 2-O′-methylation (box C/D) and pseudouridylation (box H/ACA) of ribosomal RNA (rRNA) and small nuclear RNAs (snRNA), respectively (Figure 1A) [1]. A number of snoRNAs also directly participate in the nucleolytic processing of rRNA precursors [5]. However, a large subclass of snoRNAs called orphan snoRNAs appears not to match any other RNA sequence [6]. Higher organisms also encode many noncanonical snoRNAs that may lack specific boxes, may be shortened or extended, or may contain both C/D and H/ACA boxes [6–9]. Although expression of many of these unusual snoRNAs is yet to be confirmed, they are likely to have alternative regulatory functions, for example, in pre-mRNA splicing [7,8] or regulation of polyadenylation site (PAS) recognition via direct snoRNA interaction with the

[1]Institute of Genetics and Biotechnology, Faculty of Biology, University of Warsaw, Pawinskiego 5a, 02-106 Warsaw, Poland
[2]School of Biosciences, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[@]Twitter: @grzechnik_pawel

*Correspondence: P.L.Grzechnik@bham.ac.uk (P. Grzechnik).

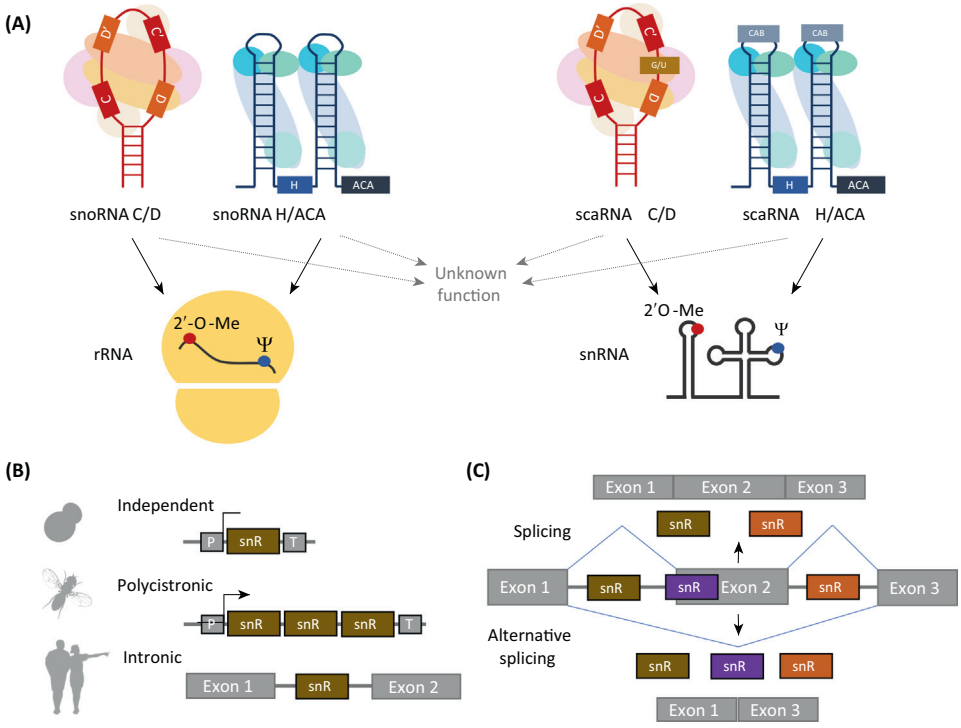

Figure 1. Classes and Genomic Organisation of Small Nucleolar RNAs (snoRNAs). (A) snoRNA classes and functions. Conserved consensus sequences (boxes) are shown as rectangles. 2′-O-Me: box C/D-dependent methylation; Ψ: box H/ACA-dependent pseudouridylation. (B) Genomic organisation of snoRNA genes changes with increasing organism complexity. (C) Alternative splicing regulates expression of snoRNA genes located within the same host molecule. snoRNA, which partially overlaps with Exon 2, is not expressed when all exons are spliced. Exon 2 skipping generates an snoRNA-containing lariat, the processing of which releases all three snoRNAs. Exon, mRNA or lncRNA exons; P, promoter; scaRNA, small Cajal body-associated RNA; T, terminator.

mRNA cleavage and polyadenylation (CPA) machinery [10]. snoRNAs may also affect cellular processes indirectly, by disturbing ribosome and snRNA formation, which as a result may control expression of protein-coding genes via splicing and translation efficiency [11,12]. Another indirect regulatory level is provided by snoRNAs that are processed to shorter microRNAs (miRNAs) [9] or Piwi-interacting RNAs (piRNAs), which act in gene silencing pathways [13].

The number of snoRNA genes increases with organism complexity. For example, in the yeast *Saccharomyces cerevisiae*, 64 transcription units encode 76 snoRNAs [14]. In humans, it is estimated that there are ~550 snoRNAs; however, *in silico* analyses predict that the total number of different types of snoRNAs may exceed 1000 [6,15]. Strikingly, genomic organisation of snoRNA genes also follows the evolutionary tree (Figure 1B). *S. cerevisiae* snoRNAs are mainly transcribed from independent genes; however, very few are organised in polycistrons (17 snoRNAs localised in 5 clusters) or are embedded in mRNA introns (8 snoRNAs) [14]. In *Drosophila melanogaster* and *Caenorhabditis elegans*, snoRNAs are less frequently transcribed as single units and are mostly organised in polycistrons [1]. In humans, snoRNA genes are predominantly located within mRNA or long noncoding RNA (lncRNA) introns and released from the host precursor by splicing [1,16]. However, snoRNA synthesis does not necessarily

correlate with the expression of mRNA of the host genes. Up to 96% of 173 transcribed snoRNA host genes in human cells frequently produce mRNA or lncRNA isoforms that are directed to degradation via nonsense-mediated RNA decay (NMD) [17]. This suggests that at least some snoRNA hosting mRNAs and lncRNAs are nonfunctional and are transcribed and spliced only to release intronic snoRNAs. Also, in the case of multi-snoRNA hosts, some snoRNA genes are located fully or partially within their exons. Many isoforms generated by alternative splicing of such transcripts become NMD substrates, allowing for selective expression of intronic snoRNAs, but not those located in exons [17]. This illustrates that snoRNA accumulation can be regulated independently from the level of the host RNA or other snoRNAs within the same transcription unit (Figure 1C).

snoRNA Maturation Pathways

snoRNA genes are usually transcribed by RNA Polymerase II (Pol II), which also synthesises mRNAs. However, in contrast to mRNAs, which are exported to the cytoplasm and translated, snoRNAs remain in the nucleus [1,18,19]. This may be facilitated by the lack of mRNA structural hallmarks: the m^7G cap bound by the cap-binding complex (CBC) and the poly(A) tail, which both enhance export to the cytoplasm. These mRNA-specific elements are removed by snoRNA maturation. Indeed, in *S. cerevisiae*, snoRNAs that retain an mRNA-like 5′ end are exported from the nucleus [20].

The general outline of snoRNA processing follows the same course for all classes in all organisms: newly synthesised pre-snoRNAs are trimmed by ribonucleases at 3′ or both 5′ and 3′ ends. As a result, 5′ and 3′ RNA regions that are not protected by cotranscriptionally associated snoRNP proteins or RNA secondary structures, are removed. In cases where the transcription start site (TSS) defines the mature 5′ end, the cap is retained on snoRNA and is post-transcriptionally modified (see below).

snoRNA 3′ End Processing in *S. cerevisiae*

The 3′ end processing, coupled with transcription termination, is one of the most important processes defining an RNA precursor as either mRNA or ncRNA. In eukaryotic cells, mRNA-specific transcription termination and 3′ end processing, mediated by the CPA complex, lead to mRNA stabilisation (Box 1), while short noncoding genes employ mechanisms which couple transcription termination with 3′-5′ degradation [21,22].

snoRNA 3′ end formation is determined by genomic organisation and employs different factors in various organisms (Figure 2, Key Figure). In *S. cerevisiae*, transcription termination of independently transcribed snoRNAs is mediated by the ncRNA-specific NNS (Nrd1-Nab3-Sen1) complex (Figure 2A), which consists of the RNA-binding proteins Nrd1 and Nab3 and the RNA:DNA helicase Sen1 [22]. The NNS complex interacts with the Pol II C-terminal domain (CTD) via the Nrd1 CTD-interacting domain (CID) and acts at a short distance (>1 kb) from the TSS [23–25]. Transcription termination is initiated by binding of Nrd1 and Nab3 to NNS-binding sites (NBS) (GUA[A/G] and UCUU[G], respectively) appearing in the nascent RNA [22,26,27]. This decreases the Pol II transcription rate downstream of NBS, which mirrors the process observed for protein-coding genes downstream of PAS. Pol II is released from the DNA by the helicase Sen1 in a mechanism that resembles Rho-dependent termination in bacteria [28]. Sen1 can both translocate along a single-stranded RNA or unwind short RNA:DNA hybrids *in vitro* [28]. Thus, *in vivo* Sen1 may use either strategy to translocate towards Pol II and release it from DNA by pulling the RNA from the polymerase catalytic centre. An analogous strategy may be employed by the 5′-3′ exonuclease Rat1 (XRN2 in humans) for protein-coding genes (Box 1) [21]. In addition to NNS, transcription termination of yeast ncRNA requires a number of mRNA-

Box 1. Transcription Termination and 3′ End Processing of Protein-Coding Genes

The mechanism of transcription termination and mRNA 3′ end processing is conserved in eukaryotic organisms and mediated by numerous factors, generally called the cleavage and polyadenylation (CPA) complex. In *S. cerevisiae*, the central core of the CPA machinery is a multi-subunit cleavage and polyadenylation factor (CPF). Recent analysis revealed that CPF consists of nuclease, poly(A) polymerase, and phosphatase modules [92]. The last module is also present in the APT (associated with Pta1) complex, which is required for transcription termination of ncRNAs [29,32,93]. Transcription termination of protein-coding genes also requires additional subunits that cooperate with CPF: cleavage factor IA containing the Pol II CTD-binding protein Pcf11, and cleavage factor IB. The human CPA complex is organised in a similar manner. The cleavage and polyadenylation specificity factor (CPSF), which comprises RNA endonuclease CPSF-73 and poly(A) polymerase PAP, is assisted by the cleavage stimulation factor (CstF) and cleavage factors I and II (CFIm and CFIIm with PCF11) [94]. CPA is recruited to the transcribing polymerase complex at the 3′ end of genes and binds sequences that constitute the polyadenylation signal (PAS) in the nascent RNA. The CPF endonuclease CPSF-73 (Ysh1 in yeast) cleaves pre-mRNA 10–30 nt from the PAS (AAUAAA in humans) and the released transcript is immediately polyadenylated at the 3′ end by PAP (Pap1 in yeast) [21,93]. The cleavage at PAS has several consequences. Pol II transcription rate significantly decreases and the Pol II complex may undergo conformational changes, which facilitate termination [95,96]. Moreover, the free 5′ end of the nascent RNA provides an entry site for 5′-3′ exonuclease XRN2 (Rat1 in yeast). As a result, nascent RNA, which is generated by Pol II at a slow pace, is rapidly digested in the 5′ to 3′ direction. XRN2/Rat1 pursues the transcribing complex and it is anticipated that the exonuclease finally drags the RNA from Pol II. This may disrupt the DNA:RNA hybrid in the active centre and release Pol II from the DNA template. The average distance of Pol II dissociation in humans is 3300 nt from the cleavage site and 160–190 nt in yeast cells [97,98].

specific CPA factors. These include the phosphatase module of the cleavage and polyadenylation factor (CPF, see Box 1), which together with Syc1 protein form the APT (associated with Pta1) complex, and the component of the cleavage factor IA, CTD-interacting protein Pcf11 [29–32]. While the exact contribution of APT to snoRNA termination is not fully understood, Pcf11 has been extensively studied. Pcf11 is recruited to the transcribing complex downstream of functional NBS clusters, which indicates that Pcf11 acts downstream of the NNS complex in the termination process [31]. One scenario assumes that Pcf11 and CPA components may employ AT-rich tracts in snoRNA terminators, thereby providing a fail-safe mechanism [26]. Alternatively, Pcf11, either in the CPA complex or independently, collaborates with NNS. The presence of Nrd1 at snoRNA terminators stimulates Pcf11 cotranscriptional recruitment, which in turn displaces Nrd1 on the Pol II CTD [31]. This may further impair Pol II transcription and facilitate Sen1-dependent termination, which requires the presence of Pcf11. This option is supported by observations that Pcf11 mutation results in transcription termination defects for snoRNAs, as well as in accumulation of RNA:DNA hybrids downstream of NNS terminators [31]. The fission yeast *Schizosaccharomyces pombe* Nrd1 homologue Seb1 functionally resembles Pcf11 and is involved in transcription termination of both coding RNAs and ncRNAs [33,34].

The NNS complex recruits 3′ end processing factors (Figure 3A). First, pre-snoRNAs are transiently oligoadenylated at the 3′ end by the TRAMP complex [35,36]. In contrast to mRNA, the adenine tail is short and used to attract exonucleases. The TRAMP poly(A) polymerase Trf4 binds Nrd1 through the Nrd1-interacting motif that mimics Pol II CTD [37]. The TRAMP complex also recruits 3′-5′ processing enzymes (the exosome and its nuclear cofactor exonuclease Rrp6 [37,38]), which exonucleolytically degrade pre-snoRNA until they encounter a formed snoRNP [39]. In fission yeast, where NNS is not present, the exosome is recruited to pre-snoRNAs via the poly(A)-binding protein Pab2 [40]. In a few cases (e.g., U3, snR40), the endoribonuclease Rnt1 (RNase III) cleaves snoRNA precursors downstream of the mature sequence, generating an entry site for exosome- and Rrp6-mediated processing [41,42].

Biogenesis of intronic snoRNAs in *S. cerevisiae* is not coupled to transcription termination, but depends on splicing and snoRNA release by the intron-debranching factor Dbr1 or Rnt1

Key Figure

Processing Strategies for Small Nucleolar RNAs (snoRNAs)

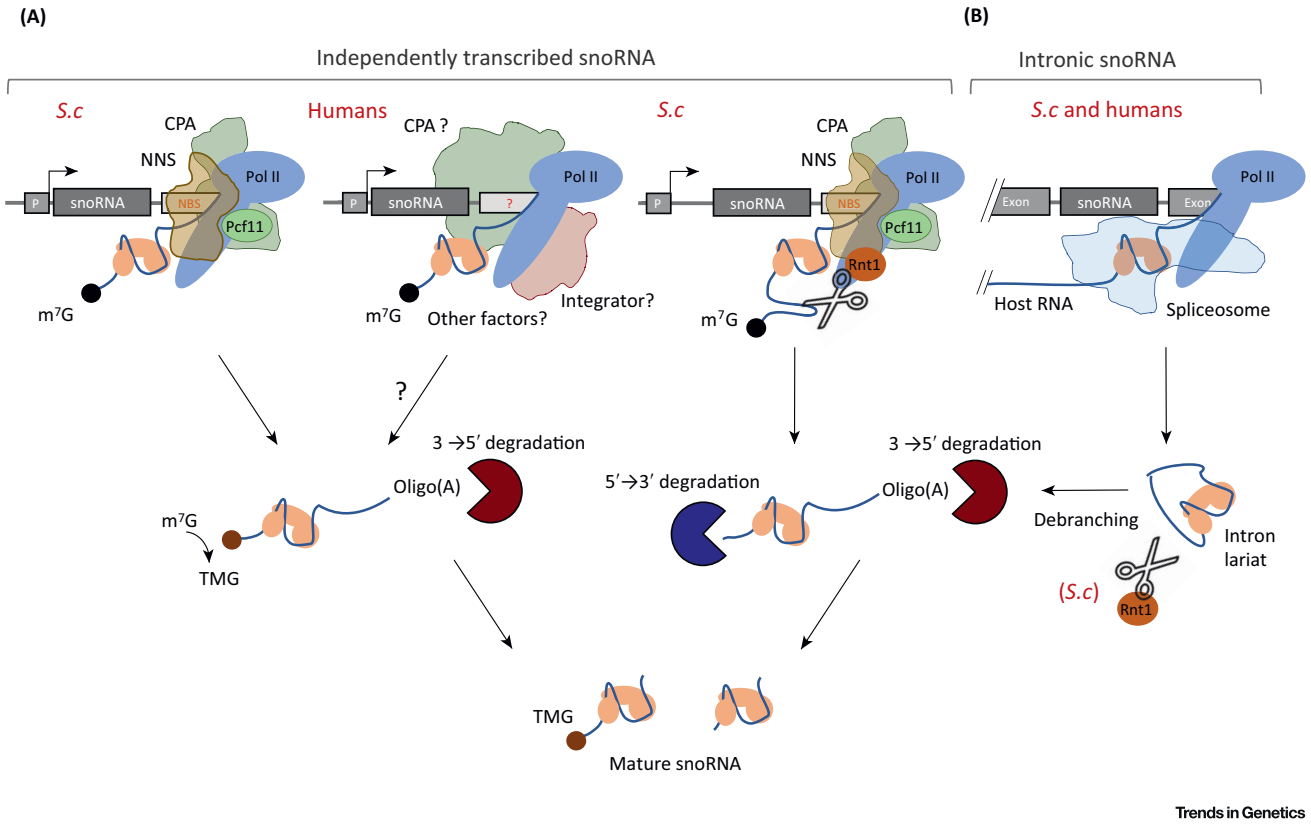

Figure 2. (A) Processing of individually transcribed snoRNAs. In *Saccharomyces cerevisiae*, snoRNAs are terminated by the Nrd1-Nab3-Sen1 (NNS) complex, assisted by some cleavage and polyadenylation factors, including Pcf11. Factors and sequences mediating transcription termination in human cells remain mostly unknown; however, the CPA factors and the Integrator complexes are possible candidates. The m^7G cap is cotranscriptionally removed by Rnt1 from the majority of box C/D pre-snoRNAs in *S. cerevisiae*, or post-transcriptionally converted to a 2,2,7-trimethylguanosine (TMG) cap for other snoRNAs. Unprotected snoRNA ends are processed by exonucleases, Rat1 (5′-3′) and the nuclear exosome (3′-5′). Exonucleases advance along RNA until they are blocked by the snoRNP structure. (B) Intronic snoRNAs are processed by exonucleases, mainly from spliced out and debranched introns. CPA, Cleavage and polyadenylation complex; NBS, NNS-binding site; P, promoter; *S.c.*, *S. cerevisiae*.

cleavage (Figure 2B) [43–45]. Some snoRNAs (U24, snR38) strictly rely on splicing, while others (U18, snR39, snR59) employ both splicing-dependent and Rnt1-mediated pathways. Following the action of Dbr1 or Rnt1, 3′ ends of pre-snoRNAs are processed exonucleolytically, but the exosome and Rrp6 recruitment mode to these precursors has not been described. It is possibly mediated by components of the NNS complex, Nrd1 and Nab3, which post-transcriptionally bind many RNAs to mark them for exosome-dependent degradation [46]. In the case of snoRNA clusters, individual snoRNAs are released from the polycistronic transcript by Rnt1-mediated cleavage inbetween each snoRNA, whereas the last snoRNA undergoes termination-dependent maturation similarly to independently transcribed snoRNAs [41,47,48].

snoRNA 3′ End Processing in Humans

In human cells only a small fraction of snoRNA genes are expressed as independent units [1,49]. Although this includes essential U3, U8, U13, and telomerase snoRNA, their biogenesis has not been extensively studied. It is plausible that these snoRNAs may employ a mechanism similar to other short ncRNAs (Figure 2A). Since the NNS complex is not present in human cells, short ncRNAs, such as snoRNA-like snRNAs (Box 2), enhancer RNAs (eRNAs), and promoter upstream transcripts (PROMPTs), are generally terminated by CPA factors with a possible involvement of the 3′ end processing Integrator complex (Box 2). The Integrator promotes endonucleolytic cleavage of the nascent snRNA and eRNA, resulting in their release from Pol II, which contributes to transcription termination [50–52]. In addition to CPA and the Integrator, termination of ncRNAs requires a protein, ARS2, which functionally resembles the yeast NNS complex and links transcription termination with 3′ end processing [53–55]. ARS2 is recruited to RNA by the CBC to form the CBCA complex, and stimulates transcription termination of snRNAs, eRNAs, and PROMPTs in the proximity of the TSS [54,55]. Although the interaction of CBCA with PHAX is essential for targeting of U3, U8, and U13 snoRNAs to Cajal bodies [56], it is not clear whether CBCA is also involved in snoRNA 3′ end formation, since depletion of ARS2 and CBC does not affect transcription termination of U8 snoRNA [54], The 3′ end processing mode of independent snoRNAs raises further questions. Their maturation may be carried out by the exosome that is recruited to precursors by the NEXT complex, which forms the CBCN complex with CBCA via its component ZC3H18 protein [53,57]. However, in the case of other short ncRNAs, NEXT is excluded from RNA precursors by PHAX, which competes with ZC3H18 for CBCA binding [58]. Alternatively, NEXT may be loaded on 3′ unprocessed snoRNAs independently of CBC via its RNA-binding protein RBM7. Such a mechanism has been observed for human intron-encoded snoRNAs [58,59]. These intronic snoRNAs are processed exonucleolytically from spliced and debranched introns or precursors released by endonucleolytic cleavages (Figures 2B and 3B) [60–63]. In both cases, RBM7 binds to long RNA 3′ extensions, downstream of mature RNAs, where it most likely recruits NEXT, thus marking these precursors for exosome-dependent processing. It is also possible that individual and intronic pre-snoRNAs targeted by RBM7/NEXT and the exosome are mainly subjected to complete degradation. Consistently, snoRNA 3′ end maturation has been reported to be mediated by deadenylases PARN and TOE1 (Figure 3B), which remove poly(A) tails added post-transcriptionally by the PAPD5/TRF4-2 component of the human TRAMP complex, and also trim redundant genome-encoded residues [64–66]. PARN and TOE1 functions generally overlap, but PARN appears to be more specific for 3′ end processing of nucleolar-localised snoRNAs and telomerase RNA, while TOE1 is engaged in 3′ end formation of scaRNAs. Both deadenylases probably compete with PAPD5 and the exosome to maintain the balance between processing and degradation [65,66].

snoRNA 5′ End Processing

Mature snoRNAs also differ from mRNAs at their 5′ ends. The m^7G cap is added cotranscriptionally to every Pol II transcript, including individually transcribed snoRNAs, when the nascent RNA is about 20–30 nt long [67]. However, pre-snoRNA 5′ end processing either removes the cap or converts the m^7G to a 2,2,7-trimethylguanosine (TMG) cap (Figure 2A). *S. cerevisiae* employs both strategies; TMG capping dominates for box H/ACA, while cap removal is more often applied for maturation of box C/D snoRNAs [20,41,68]. In yeast, the majority of box C/D snoRNAs are transcribed with a long (150–200 nt) 5′ extension, which forms a stem loop recognised and cleaved by the endonuclease Rnt1. The resulting unprotected 5′ end is attacked by the 5′-3′ exonucleases Rat1 and Xrn1 [41,68,69]. Both enzymes degrade precursors from the 5′ end until they are blocked by the snoRNP structure.

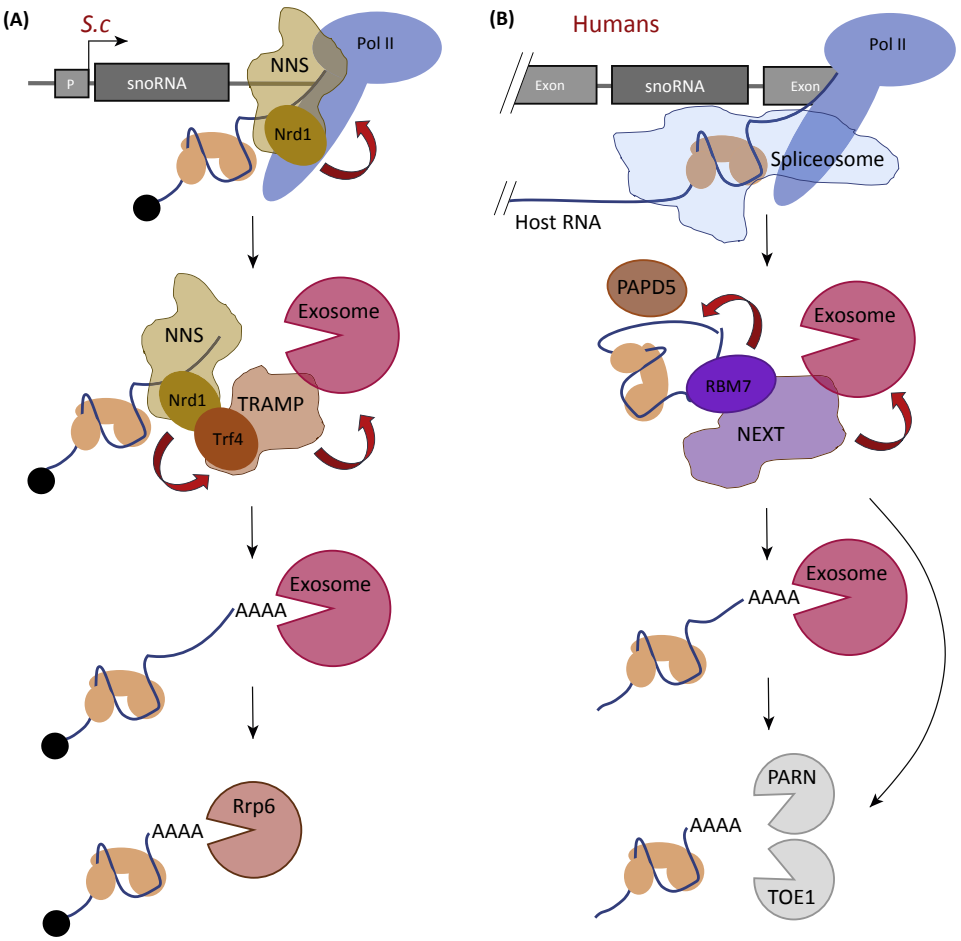

Figure 3. Transition from Transcription to Small Nucleolar RNA (snoRNA) 3′ End Processing. (A) The Nrd1-Nab3-Sen1 (NNS) component Nrd1 mediates handover from termination to 3′ end processing in *Saccharomyces cerevisiae*. The NNS complex initiates transcription termination by interacting via Nrd1 with both nascent RNA and the Pol II C-terminal domain (CTD). Pre-snoRNA released from the polymerase complex is bound by NNS. Nrd1, through interaction with the poly(A) polymerase Trf4, recruits the TRAMP complex, which in turn recruits the nuclear exosome. The exosome removes 3′ extensions in rounds of subsequent oligoadenylation and exonucleolytic trimming. The last nucleotides are digested by the exosome cofactor exonuclease Rrp6. (B) The exosome is recruited to intron-encoded snoRNA via the NEXT component RBM7. Trimming of the 3′ end is most likely performed by a combined action of the poly (A) polymerase PAPD5, exosome, and deadenylases PARN and TOE1, while last extended nucleotides are removed exclusively by deadenylases. *S.c.*, *S. cerevisiae*.

Rnt1 interacts with the NNS component Sen1, which suggests a crosstalk between 5′ and 3′ end processing [70]. Indeed, despite its activity at snoRNA 5′ ends, Rnt1 is cotranscriptionally recruited over their NNS terminators at the 3′ ends [20]. Such a strategy allows for snoRNP formation before the precursor is cleaved and so prevents premature 5′-3′ trimming, which may lead to complete degradation of pre-snoRNA. Removal of the cap structure is required for downstream processes in box C/D snoRNA synthesis. The 5′ unprocessed box C/D pre-snoRNAs have short, oligoadenylated unprocessed 3′ ends and become exported from the nucleus [20].

Box 2. Maturation of snRNAs

snRNAs are regarded as the closest RNA family to snoRNAs as they are similarly stable, localised in the nucleus, and share some processing strategies. Spliceosomal U1, U2, U4, and U5 snRNAs are synthesised by Pol II, unlike U6, which is a Pol III transcript and its biogenesis follows a separate pathway. In all eukaryotes snRNAs are transcribed from independent genes and are not processed at 5′ ends and therefore their m^7G caps are converted to the TMG cap by Tgs1 [71]. In turn, the snRNA 3′ end maturation pathway depends on the organism. In *S. cerevisiae* snRNAs are terminated by the NNS complex and processed at the 3′ end by the exosome and Rrp6 exonuclease, which resembles the mechanism for independently transcribed snoRNAs [22]. Alternatively, as is the case for U3 snoRNA, snRNAs are exonucleolytically processed from precursors generated by Rnt1 endonucleolytic cleavage downstream of their mature ends and protected from degradation by binding of Lhp1, the homologue of human La protein [50]. The 3′ end processing is performed in the nucleus, but snRNA cap hypermethylation may also occur in the nucleolus, as U1 is retained in this structure in the absence of Tgs1 [71]. Biogenesis of metazoan snRNAs is completely different. SnRNA synthesis requires an snRNA-specific promoter, a conserved box in the 3′ end region, and a large, 14 subunit Integrator complex [50,51]. It seems that CPA factors that mediate transcription termination of human snRNAs have little, if any, role in their 3′ end processing. Instead, this pathway utilises the Integrator complex that is recruited to snRNA genes by the Pol II CTD. Pre-snRNAs are cleaved at the 3′ box by the Integrator subunit Int11, a paralogue of the CPA component CPSF-73 [99]. The release of primary snRNA transcripts may promote Pol II transcription termination, as was shown for eRNAs [52]. Cleaved snRNA intermediates are then exported from the nucleus to the cytoplasm via the CBC-bound PHAX adaptor and the CRM1/RanGTP complex. The next steps that comprise exonucleolytic trimming by an unknown nuclease, survival motor neuron (SMN), and protein arginine methyltransferase 5 (PRMT5), facilitates snRNP assembly with the core Sm proteins and cap hypermethylation, which takes place in the cytoplasm. The assembled snRNPs are imported back to the nucleus by TMG cap-bound snurportin-1 (SPN1) and importin β that probably interacts with the SMN complex [100]. Finally, snRNAs are pseudouridylated and 2′-O-methylated by scaRNAs in Cajal bodies.

For the majority of yeast box H/ACA snoRNAs, the m^7G cap is converted to a TMG cap by a nucleolar trimethylguanosine synthetase Tgs1 [71]. Deletion of the *TGS1* gene in *S. cerevisiae* has no impact on snoRNA maturation, which indicates that cap hypermethylation does not contribute to the crosstalk between both snoRNA ends [20,68]. The m^7G cap conversion to a TMG cap also occurs in higher organisms; however, the knowledge about this process in snoRNA biogenesis is very limited. Tgs1 knockdown in human cells results in disruption of Cajal bodies, but it is not clear if this is directly associated with snoRNA or snRNA synthesis, although some steps of their formation take place in these structures [72]. Equally little is known about the enzymes involved in 5′ end trimming of vertebrate intronic snoRNAs, except that it is performed by unspecified 5′-3′ exonuclease(s) [60,61].

Why snoRNAs are generally deprived of m^7G caps by either genomic organisation or 5′ end processing is still to be explored. The m^7G cap, an mRNA hallmark, may somehow counteract processes that lead to the synthesis of box C/D snoRNPs. Indeed, in yeast cells lacking Rnt1, box C/D snoRNAs with capped 5′ extensions are mislocalised to the cytoplasm [20]. Consequently, their activity in rRNA maturation is compromised, leading to hypomodified rRNA subunits [20]. This in turn may impact ribosome biogenesis and regulation of its function for optimal translation [73]. However, there might also be other aspects related to the status of snoRNA 5′ ends. Since the CBC complex promotes RNA degradation in the nucleus by Rat1/Xrn1 and the exosome/Rrp6 [54,55,74], the persistent presence of CBC on mature m^7G capped snoRNAs may decrease their stability.

Different cellular fates of Pol II-generated snoRNA and mRNA transcripts are determined by their distinct maturation pathways. Signals in the nascent transcripts or the sequence of processing events classify the newly synthesised molecule as mRNA or snoRNA.

Synthesis of Unusual snoRNA Classes

snoRNAs whose sequences differ from classical box C/D or box H/ACA layout may employ specific processing strategies. For example, the snoRNA-like component of the human

telomerase (hTR) 3′ region folds into a box H/ACA structure [75]. The 3′ hairpin terminal stem loop in this region contains an additional BIO motif, which is required for hTR 3′ processing and RNP assembly [76]. The BIO motif is also present in some AluACA RNAs, a subclass of human box H/ACA RNPs that are processed from intronic Alu sequences [77,78]. AluACA RNAs contain unusually short or long 5′ hairpins that cannot support optimal snoRNP formation required for snoRNA stability. As a result, these RNAs are very rapidly degraded. However, AluACA RNAs, which contain the BIO motif, are stabilised by assembly into snoRNPs and accumulate in Cajal bodies [77].

In human cells, snoRNAs may also exist as hybrids with lncRNAs. Two classes of such hybrids are directly associated with Prader-Willi syndrome (PWS) [7,8] (Box 3). PWS results from the loss of expression of genes from the q11–13 region of the paternally inherited chromosome 15, which contains the *snurf-snrp* gene, and a downstream noncoding region considered primarily as a snoRNA host gene [79]. The PWS-associated minimal 108 kb deletion of paternal chromosome 15 includes snoRNA SNORD109A and the SNORD116 cluster of 29 similar snoRNAs [80,81]. Some of these snoRNAs undergo unusual processing to form chromatin-associated sno-lncRNAs (snoRNA-related long noncoding RNA) and *SPA* lncRNAs (5′ snoRNA capped and polyadenylated lncRNA) (Figure 4) [7,8]. The ends of five 1–3 kb long box C/D sno-lncRNAs correspond to intronic snoRNAs imbedded in the same intron. Thus, exonucleolytic degradation of the excised lariat is blocked by assembled snoRNPs that define mature 5′ and 3′ sno-lncRNA ends (Figure 4A) [7]. *SPA* lncRNAs are much longer (34 kb and 16 kb) than sno-lncRNAs, contain 5′ terminal box C/D snoRNPs, and are polyadenylated at the 3′ ends [8]. Their synthesis is splicing-independent and regulated by PAS-dependent transcriptional termination of the upstream *snurf-snrp* gene (Figure 4B). Pol II pausing at the proximal *snurf-snrp* weak PAS is inefficient and permits polymerase to continue transcription. Still, the nascent RNA is cleaved by the CPA complex endonuclease at this PAS and the resulting free 5′ end is cotranscriptionally digested by the 5′-3′ exonuclease XRN2. The enzyme moves on until it is physically blocked by the snoRNP structure downstream of the PAS, thus defining the 5′ end of *SPA* lncRNA. Pol II proceeds then to the next PAS, where the pre-*SPA* lncRNA is cleaved, polyadenylated, and released. The second downstream *SPA* lncRNA is synthetised in a similar manner by the still elongating Pol II complex. Both sno-lncRNAs and *SPA* lncRNAs are not released to the nucleoplasm, but are retained in the chromatin fraction [7,8]. The internal RNA sequence, which is protected from degradation by the terminal snoRNP and 3′ poly(A) tail for *SPA*, is used to sequester splicing factors to fine-tune alternative splicing and thus control gene expression (Box 3).

SPA lncRNAs are also generated in other genomic localisations. An interplay between *NOP56* mRNA and *NOP56*-derived cytoplasm-located *SPA* (*cSPA*) lncRNA acts as a major switch controlling snoRNA levels in human cells [82]. The *NOP56* gene, which encodes essential C/D core protein, contains several intron-embedded snoRNAs, including an orphan box C/D *snoRD86*. A successful *snoRD86* assembly into a snoRNP directs alternative splicing of *NOP56* mRNA, which generates *snoRD86-cSPA* instead of functional mRNA. This decreases the cellular level of NOP56 protein, which in turn suppresses box C/D snoRNP formation. Inhibition of *snoRD86* maturation allows for correct splicing of *NOP56* mRNA, its translation, and thus box C/D snoRNP synthesis [82].

Unspecified RNases can further process mammalian snoRNAs to shorter (17–39 nt long) snoRNA-derived RNA (sdRNA) species [83]. A number of sdRNA exhibit miRNA or piRNA qualities [13,84,85]. Biogenesis of a subset of sno-miRNAs requires some components of the canonical miRNA pathway, mainly Dicer for box H/ACA-derived miRNA, or Dicer and DGCR8

Box 3. snoRNA-Related Diseases

Alterations in snoRNA expression and functionality may have severe implications for human health. Many snoRNAs are linked with carcinogenesis; however, the molecular basis of this association is still obscure. For example, the locus expressing SNORD50A-SNORD50B is lost in 10–40% of common tumours, including prostate, lung, liver, and skin cancer [101]. Although both snoRNAs guide methylation of the large ribosome subunit, they also directly bind and inhibit Ras GTPases and thus act as tumour suppressors [101]. SNORD50A also regulates expression of genes involved in proliferation and apoptosis by affecting mRNA poly(A) site selection via interaction with the CPA component Fip1 [10]. Moreover, high levels of box C/D snoRNAs are associated with leukaemia and its enhanced self-renewal [102]. The excess of box C/D snoRNAs may increase rRNA methylation, which in turn may reduce translation fidelity and thus deregulate gene expression.

Deletion of SNORD109A and the snoRNA genes of the SNORD116 cluster on the paternally inherited chromosome 15 is associated with Prader-Willi syndrome (PWS) [80,81]. PWS is a genetic disorder with a prevalence of 1 in 10 000–25 000 people [79]. PWS patients display a broad pathological spectrum, including morbid obesity, dysmorphic changes (e.g., characteristic facial features, hypogonadia, and infertility), behavioural actions (temper tantrums, stubbornness, compulsive behaviours), and mild intellectual disabilities. Despite the clear connection between PWS and snoRNA deficiency, molecular mechanisms underlying this disease are yet to be uncovered. snoRNAs from the PWS-related region belong to the orphan snoRNA class. SNORD115 may directly bind and regulate alternative splicing of a few mRNAs, including serotonin receptor *5HT2CR* [103,104]. Moreover, chromatin-associated PWS snoRNA–lncRNA hybrids, called sno-lncRNAs and *SPA* lncRNAs (see main text), sequester RNA-binding proteins FOX2, hnRNP M, and TD43 [7,8]. Interestingly, spliced exons of the snoRNA host ncRNA from the PWS region also accumulate close to their transcription sites and may also act as molecular sponges [105]. Alternatively, all chromatin-associated ncRNA may be used as 'chaperones' to maintain higher-order chromatin structure [106,107]. Such three-dimensional chromatin organisation may play crucial roles in the regulation of gene expression [108]. Overall, deletions in the PWS region, although viable, may lead to minor deregulation of splicing, which, if they accumulate during development, may manifest as PWS [7,8].

for miRNAs processed from both snoRNA classes [86]. Several sno-miRNAs act in RNA silencing pathways by associating with Argonaute proteins [83]. sno-derived piRNAs may function in an RNA quality mechanism, as is the case for piR30840 that forms a complex with Ago4, Piwil4, and TRAMP to target pre-mRNA substrate for exosome-mediated degradation [13].

snoRNA: A Model ncRNA or an Exception?

Remarkable stability is a distinctive feature of snoRNAs in comparison with other transcripts generated by Pol II. This is ensured by the compact structure of snoRNAs and their association with snoRNP proteins. Even so, biogenesis of snoRNAs and some ncRNAs are similar to a certain extent. This is most apparent in yeast, where several ncRNAs, including snoRNAs, snRNAs, and unstable transcripts such as cryptic unstable transcripts (CUTs) are terminated and processed by the same machinery, with the NNS complex playing a pivotal role [22]. It appears that ncRNA degradation or processing by the NNS-recruited exosome are equivalent processes, the outcome of which is dictated by the concomitant RNP assembly. RNAs, which form highly organised secondary structures, like snoRNAs and snRNAs, are protected during nucleolytic trimming by forming stable RNP structures that halt advancing exonucleases, while other ncRNAs are fully degraded.

In higher eukaryotes, the majority of transcripts synthesised by Pol II are probably terminated by CPA factors. Here, processing strategies of snoRNAs and other ncRNAs are in most cases class-specific, however, they may share some common features. For example, many ncRNAs, including snRNAs, eRNAs, and unstable PROMPT RNAs utilise the exosome recruited by the NEXT/CBCN complex [87]. These factors may contribute to the processing of intronic snoRNAs, but they are more likely to act in their quality control and degradation [59]. Synthesis of other ncRNAs may be more dissimilar. Circular RNAs, circRNAs, and ciRNAs are generated by back-splicing or inefficient debranching, while MALAT1 and NEAT1 ncRNA contain tRNA-like

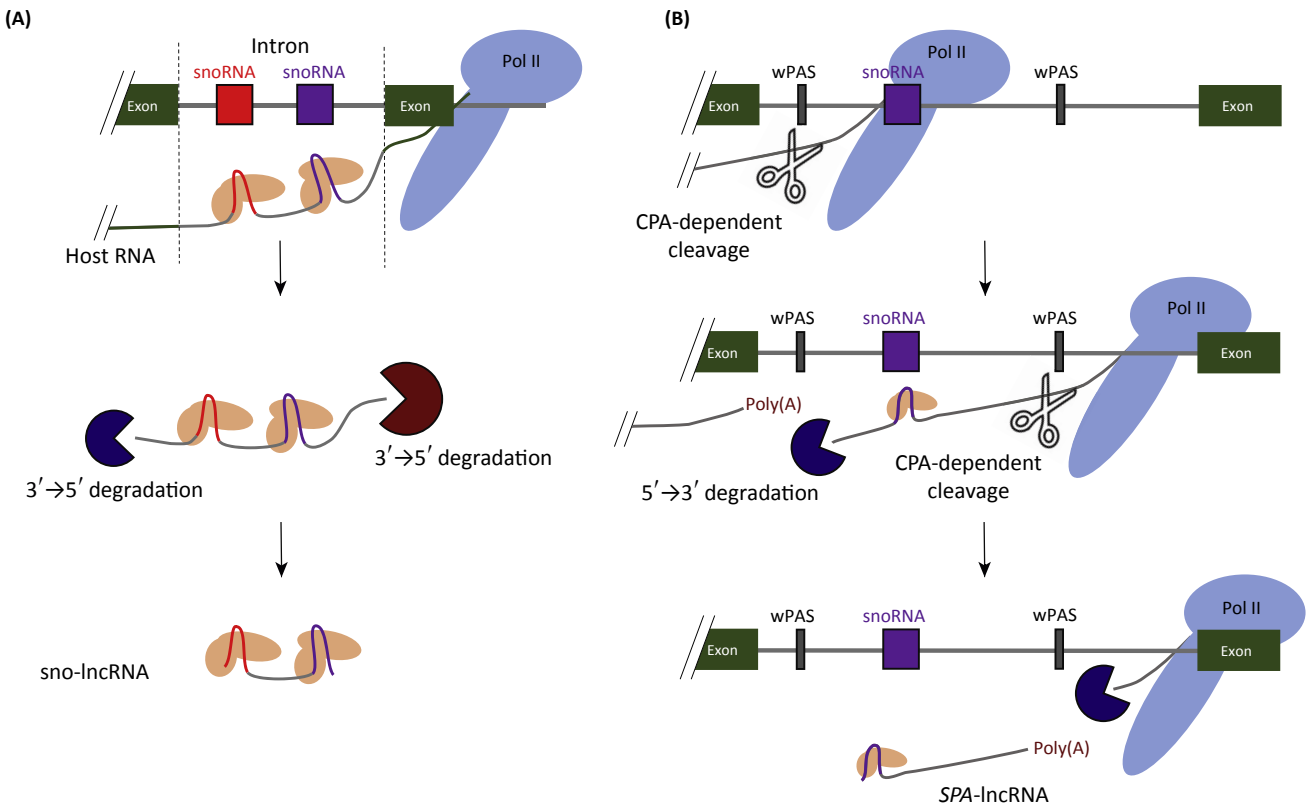

Figure 4. Unusual Processing of Noncoding RNAs (ncRNAs) Transcribed from the Prader-Willi Syndrome (PWS) Locus. (A) Biogenesis of small nucleolar RNA (snoRNA)-related long noncoding RNA (lncRNA) (sno-lncRNAs). snoRNP structures in the intron protect the RNA fragment located inbetween snoRNA. (B) Transcription and processing of 5′ snoRNA capped and polyadenylated (*SPA*) lncRNAs. Cotranscriptionally formed snoRNP blocks the advancing 5′-3′ exonuclease and thus affects transcription termination. wPAS, Weak poly(A) site.

structures and are processed by tRNA endonucleases RNase P and Z [87–89]. Finally, many lncRNAs are formed in a fashion similar to mRNAs. However, processing of ncRNAs that are capped by snoRNA structures (e.g., sno-lncRNAs and *SPAs*) or contain snoRNA-related motifs (e.g., telomerase RNA) may employ common elements with the snoRNA maturation machinery [64–66].

Differences between snoRNA and mRNA are unsurprising since their maturation is generally governed by distinct mechanisms. In yeast, the common aspects include participation of some CPA factors and the NNS complex in transcription termination of both RNA classes. The major components of the termination machinery for protein-coding genes (e.g., CTD-binding protein Pcf11) cooperate with NNS to enhance Sen1-dependent termination of snoRNA genes. However, the function of NNS in termination at protein-coding genes is either secondary or is limited to specific cases [90]. Also, initial polyadenylation of snoRNA precursors has a different role than for mRNA, as it serves to stimulate exosome-mediated 3′ end processing and is absent in mature snoRNAs [22]. Finally, involvement of the key snoRNA processing endo- or exonucleases is limited to Rnt1 activity in fail-safe termination of a few mRNAs [90]. In vertebrates, biogenesis of mRNA and snoRNA is more disparate. Here, snoRNA synthesis

generally depends on splicing and is hardly considered an independent process, although some snoRNA hosts are transcribed only to yield snoRNAs. mRNA transcription termination and 3′ end formation factors are unlikely to have any role in biogenesis of intronic snoRNAs. The activity of PARN and TOE1 deadenylases in snoRNAs 3′ end trimming also represents a very distinctive step [64–66].

Concluding Remarks

In the past decade, it became apparent that up to 90% of the eukaryotic genome is transcribed, but protein-coding genes comprise only a small fraction (2–3%), with the majority of the transcriptome consisting of ncRNAs [91]. ncRNAs represent different structural and functional classes and are regulated by a broad range of processing and decay mechanisms. Although transcribed by the same RNA polymerase II, snoRNAs and mRNAs have specific features that mark them for distinct processing pathways and specific cellular fate. These hallmarks, depending on the organism, include cis-acting elements that control their expression via specialised protein complexes (e.g., the NNS complex in yeast), or the status of their 5′ ends, which determines their processing mode and cellular destination via cap-interacting proteins.

Further research is needed to investigate the interplay between snoRNAs and host mRNA/lncRNA synthesis in vertebrates (see Outstanding Questions). How often and in which physiological conditions does alternative splicing generate snoRNA–mRNA hybrids and how does this process regulate expression of the host gene? Moreover, since only a fraction of host mRNAs and lncRNAs seem to be used as vessels for snoRNA synthesis and are then degraded after splicing, questions arise concerning the nature of transcriptional switches that decide the fate of the intronic snoRNAs and the host RNA and how synthesis of the two is coordinated. Finally, 3′ end formation of a few essential, independently transcribed snoRNAs in human cells still remains unclear.

Outstanding Questions

What is the exact role of the mRNA-specific transcription termination and 3′ end formation machinery in snoRNA biogenesis?

Which factors terminate transcription and process 3′ ends of independently transcribed snoRNAs in human cells?

How important is snoRNA 5′ end processing and why is removal of the m^7G cap important for maturation of box C/D snoRNAs? What is the exact role of snoRNA cap hypermethylation? How has the demand for decapped snoRNA contributed to the evolutionary drive in higher organisms, which has arranged snoRNA genes in clusters or placed snoRNA in intronic locations?

How many more unusual snoRNAs, either encoded in genomic sequences or generated by alternative processing, are yet to be discovered?

Considering that box C/D and box H/ACA snoRNAs are structurally distinct and, at least in *S. cerevisiae*, have unrelated processing pathways, should these two classes be considered separately in transcriptional analyses?

Acknowledgments

We thank D. Cunningham, K. Winczura, T.H. Jensen, S. Lykke-Andersen, E. Bertrand, and C. Verheggen for critical reading of the manuscript. J.K. was funded by Polish-Swiss Research Programme (PSPB-183/2010). P.G. was supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (200473/Z/16/Z).

References

1. Dieci, G. *et al.* (2009) Eukaryotic snoRNAs: a paradigm for gene expression flexibility. *Genomics* 94, 83–88
2. Darzacq, X. *et al.* (2002) Cajal body-specific small nuclear RNAs: a novel class of 2′-O-methylation and pseudouridylation guide RNAs. *EMBO J.* 21, 2746–2756
3. Marnef, A. *et al.* (2014) Targeting vertebrate intron-encoded box C/D 2′-O-methylation guide RNAs into the Cajal body. *Nucleic Acids Res.* 42, 6616–6629
4. Massenet, S. *et al.* (2017) Assembly and trafficking of box C/D and H/ACA snoRNPs. *RNA Biol.* 14, 680–692
5. Watkins, N.J. and Bohnsack, M.T. (2012) The box C/D and H/ACA snoRNPs: key players in the modification, processing and the dynamic folding of ribosomal RNA. *Wiley Interdiscip. Rev. RNA* 3, 397–414
6. Jorjani, H. *et al.* (2016) An updated human snoRNAome. *Nucleic Acids Res.* 44, 5068–5082
7. Yin, Q.F. *et al.* (2012) Long noncoding RNAs with snoRNA ends. *Mol. Cell* 48, 219–230
8. Wu, H. *et al.* (2016) Unusual processing generates SPA lncRNAs that sequester multiple RNA binding proteins. *Mol. Cell* 64, 534–548
9. Falaleeva, M. and Stamm, S. (2013) Processing of snoRNAs as a new source of regulatory non-coding RNAs: snoRNA fragments form a new class of functional RNAs. *Bioessays* 35, 46–54
10. Huang, C. *et al.* (2017) A snoRNA modulates mRNA 3′ end processing and regulates the expression of a subset of mRNAs. *Nucleic Acids Res.* 45, 8647–8660
11. Ruggero, D. and Pandolfi, P.P. (2003) Does the ribosome translate cancer? *Nat. Rev. Cancer* 3, 179–192
12. Karijolich, J. and Yu, Y.T. (2010) Spliceosomal snRNA modifications and their function. *RNA Biol.* 7, 192–204
13. Zhong, F. *et al.* (2015) A snoRNA-derived piRNA interacts with human interleukin-4 pre-mRNA and induces its decay in nuclear exosomes. *Nucleic Acids Res.* 43, 10474–10491
14. Piekna-Przybylska, D. *et al.* (2007) New bioinformatic tools for analysis of nucleotide modifications in eukaryotic rRNA. *RNA* 13, 305–312
15. Hubbard, T.J. *et al.* (2009) Ensembl 2009. *Nucleic Acids Res.* 37, D690–D697
16. Makarova, J.A. and Kramerov, D.A. (2011) SNOntology: myriads of novel snoRNAs or just a mirage? *BMC Genomics* 12, 543
17. Lykke-Andersen, S. *et al.* (2014) Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. *Genes Dev.* 28, 2498–2517

18. Huang, Y. and Carmichael, G.G. (1996) Role of polyadenylation in nucleocytoplasmic transport of mRNA. *Mol. Cell. Biol.* 16, 1534–1542
19. Dunn, E.F. *et al.* (2005) Yeast poly(A)-binding protein, Pab1, and PAN, a poly(A) nuclease complex recruited by Pab1, connect mRNA biogenesis to export. *Genes Dev.* 19, 90–103
20. Grzechnik, P. *et al.* (2018) Nuclear fate of yeast snoRNA is determined by co-transcriptional Rnt1 cleavage. *Nat. Commun.* 9, 1783
21. Proudfoot, N.J. (2016) Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. *Science* 352, aad9926
22. Porrua, O. *et al.* (2016) Transcription termination: variations on common themes. *Trends Genet.* 32, 508–522
23. Steinmetz, E.J. *et al.* (2006) Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell* 24, 735–746
24. Gudipati, R.K. *et al.* (2008) Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice. *Nat. Struct. Mol. Biol.* 15, 786–794
25. Arigo, J.T. *et al.* (2006) Regulation of yeast NRD1 expression by premature transcription termination. *Mol. Cell* 21, 641–651
26. Porrua, O. *et al.* (2012) In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination. *EMBO J.* 31, 3935–3948
27. Schulz, D. *et al.* (2013) Transcriptome surveillance by selective termination of noncoding RNA synthesis. *Cell* 155, 1075–1087
28. Porrua, O. and Libri, D. (2013) A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast. *Nat. Struct. Mol. Biol.* 20, 884–891
29. Nedea, E. *et al.* (2003) Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3′-ends. *J. Biol. Chem.* 278, 33000–33010
30. Kim, M. *et al.* (2006) Distinct pathways for snoRNA and mRNA termination. *Mol. Cell* 24, 723–734
31. Grzechnik, P. *et al.* (2015) Pcf11 orchestrates transcription termination pathways in yeast. *Genes Dev.* 29, 849–861
32. Lidschreiber, M. *et al.* (2018) The APT complex is involved in non-coding RNA transcription and is distinct from CPF. *Nucleic Acids Res.* 46, 11528–11538
33. Lemay, J.F. *et al.* (2016) The Nrd1-like protein Seb1 coordinates cotranscriptional 3′ end processing and polyadenylation site selection. *Genes Dev.* 30, 1558–1572
34. Wittmann, S. *et al.* (2017) The conserved protein Seb1 drives transcription termination by binding RNA polymerase II and nascent RNA. *Nat. Commun.* 8, 14861
35. LaCava, J. *et al.* (2005) RNA degradation by the exosome is promoted by a nuclear polyadenylation complex. *Cell* 121, 713–724
36. Schmidt, K. and Butler, J.S. (2013) Nuclear RNA surveillance: role of TRAMP in controlling exosome specificity. *Wiley Interdiscip. Rev. RNA* 4, 217–231
37. Tudek, A. *et al.* (2014) Molecular basis for coordinating transcription termination with noncoding RNA degradation. *Mol. Cell* 55, 467–481
38. Vasiljeva, L. and Buratowski, S. (2006) Nrd1 interacts with the nuclear exosome for 3′ processing of RNA polymerase II transcripts. *Mol. Cell* 21, 239–248
39. Grzechnik, P. and Kufel, J. (2008) Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast. *Mol. Cell* 32, 247–258
40. Lemay, J.F. *et al.* (2010) The nuclear poly(A)-binding protein interacts with the exosome to promote synthesis of noncoding small nucleolar RNAs. *Mol. Cell* 37, 34–45
41. Chanfreau, G. *et al.* (1998) Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism. *J. Mol. Biol.* 284, 975–988
42. Kufel, J. *et al.* (2000) Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol. Cell. Biol.* 20, 5415–5424
43. Ooi, S.L. *et al.* (1998) Intronic snoRNA biosynthesis in *Saccharomyces cerevisiae* depends on the lariat-debranching enzyme: intron length effects and activity of a precursor snoRNA. *RNA* 4, 1096–1110
44. Ghazal, G. *et al.* (2005) Genome-wide prediction and analysis of yeast RNase III-dependent snoRNA processing signals. *Mol. Cell. Biol.* 25, 2981–2994
45. Villa, T. *et al.* (1998) Processing of the intron-encoded U18 small nucleolar RNA in the yeast *Saccharomyces cerevisiae* relies on both exo- and endonucleolytic activities. *Mol. Cell. Biol.* 18, 3376–3383
46. Wlotzka, W. *et al.* (2011) The nuclear RNA polymerase II surveillance system targets polymerase III transcripts. *EMBO J.* 30, 1790–1803
47. Chanfreau, G. *et al.* (1998) Processing of a dicistronic small nucleolar RNA precursor by the RNA endonuclease Rnt1. *EMBO J.* 17, 3726–3737
48. Qu, L.H. *et al.* (1999) Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast. *Mol. Cell. Biol.* 19, 1144–1158
49. Li, T. *et al.* (2010) Identification and characterization of human snoRNA core promoters. *Genomics* 96, 50–56
50. Peart, N. *et al.* (2013) Non-mRNA 3′ end formation: how the other half lives. *Wiley Interdiscip. Rev. RNA* 4, 491–506
51. Guiro, J. and Murphy, S. (2017) Regulation of expression of human RNA polymerase II-transcribed snRNA genes. *Open Biol.* 7, 170073
52. Lai, F. *et al.* (2015) Integrator mediates the biogenesis of enhancer RNAs. *Nature* 525, 399–403
53. Iasillo, C. *et al.* (2017) ARS2 is a general suppressor of pervasive transcription. *Nucleic Acids Res.* 45, 10229–10241
54. Andersen, P.R. *et al.* (2013) The human cap-binding complex is functionally connected to the nuclear RNA exosome. *Nat. Struct. Mol. Biol.* 20, 1367–1376
55. Hallais, M. *et al.* (2013) CBC-ARS2 stimulates 3′-end maturation of multiple RNA families and favors cap-proximal processing. *Nat. Struct. Mol. Biol.* 20, 1358–1366
56. Boulon, S. *et al.* (2004) PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol. Cell* 16, 777–787
57. Winczura, K. *et al.* (2018) Characterizing ZC3H18, a multidomain protein at the interface of RNA production and destruction decisions. *Cell Rep.* 22, 44–58
58. Giacometti, S. *et al.* (2017) Mutually exclusive CBC-containing complexes contribute to RNA fate. *Cell Rep.* 18, 2635–2650
59. Lubas, M. *et al.* (2015) The human nuclear exosome targeting complex is loaded onto newly synthesized RNA to direct early ribonucleolysis. *Cell Rep.* 10, 178–192
60. Kiss, T. and Filipowicz, W. (1995) Exonucleolytic processing of small nucleolar RNAs from pre-mRNA introns. *Genes Dev.* 9, 1411–1424
61. Caffarelli, E. *et al.* (1996) Processing of the intron-encoded U16 and U18 snoRNAs: the conserved C and D boxes control both the processing reaction and the stability of the mature snoRNA. *EMBO J.* 15, 1121–1131
62. Hirose, T. *et al.* (2003) Splicing-dependent and -independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. *Mol. Cell* 12, 113–123
63. Szczepinska, T. *et al.* (2015) DIS3 shapes the RNA polymerase II transcriptome in humans by degrading a variety of unwanted transcripts. *Genome Res.* 25, 1622–1633
64. Berndt, H. *et al.* (2012) Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA* 18, 958–972

65. Nguyen, D. *et al.* (2015) A polyadenylation-dependent 3′ end maturation pathway is required for the synthesis of the human telomerase RNA. *Cell Rep.* 13, 2244–2257
66. Son, A. *et al.* (2018) PARN and TOE1 constitute a 3′ end maturation module for nuclear non-coding RNAs. *Cell Rep.* 23, 888–898
67. Rasmussen, E.B. and Lis, J.T. (1993) In vivo transcriptional pausing and cap formation on three *Drosophila* heat shock genes. *Proc. Natl. Acad. Sci. U. S. A.* 90, 7923–7927
68. Lee, C.Y. *et al.* (2003) The roles of endonucleolytic cleavage and exonucleolytic digestion in the 5′-end processing of *S. cerevisiae* box C/D snoRNAs. *RNA* 9, 1362–1370
69. Petfalski, E. *et al.* (1998) Processing of the precursors to small nucleolar RNAs and rRNAs requires common components. *Mol. Cell. Biol.* 18, 1181–1189
70. Finkel, J.S. *et al.* (2010) Sen1p performs two genetically separable functions in transcription and processing of U5 small nuclear RNA in *Saccharomyces cerevisiae*. *Genetics* 184, 107–118
71. Mouaikel, J. *et al.* (2002) Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell* 9, 891–901
72. Lemm, I. *et al.* (2006) Ongoing U snRNP biogenesis is required for the integrity of Cajal bodies. *Mol. Biol. Cell* 17, 3221–3231
73. Sloan, K.E. *et al.* (2017) Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* 14, 1138–1152
74. Das, B. *et al.* (2003) Degradation of normal mRNA in the nucleus of *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 23, 5502–5515
75. Nguyen, T.H.D. *et al.* (2018) Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature* 557, 190–195
76. Jady, B.E. *et al.* (2004) Human telomerase RNA and box H/ACA scaRNAs share a common Cajal body-specific localization signal. *J. Cell Biol.* 164, 647–652
77. Ketele, A. *et al.* (2016) Human intron-encoded AluACA RNAs and telomerase RNA share a common element promoting RNA accumulation. *RNA Biol.* 13, 1274–1285
78. Jady, B.E. *et al.* (2012) Human intron-encoded Alu RNAs are processed and packaged into Wdr79-associated nucleoplasmic box H/ACA RNPs. *Genes Dev.* 26, 1897–1910
79. Angulo, M.A. *et al.* (2015) Prader-Willi syndrome: a review of clinical, genetic, and endocrine findings. *J. Endocrinol. Invest.* 38, 1249–1263
80. Duker, A.L. *et al.* (2010) Paternally inherited microdeletion at 15q11.2 confirms a significant role for the SNORD116C/D box snoRNA cluster in Prader-Willi syndrome. *Eur. J. Hum. Genet.* 18, 1196–1201
81. Sahoo, T. *et al.* (2008) Prader-Willi phenotype caused by paternal deficiency for the HBII-85C/D box small nucleolar RNA cluster. *Nat. Genet.* 40, 719–721
82. Lykke-Andersen, S. *et al.* (2018) Box C/D snoRNP autoregulation by a cis-acting snoRNA in the NOP56 pre-mRNA. *Mol. Cell* 72, 99–111
83. Dupuis-Sandoval, F. *et al.* (2015) The emerging landscape of small nucleolar RNAs in cell biology. *Wiley Interdiscip. Rev. RNA* 6, 381–397
84. Ender, C. *et al.* (2008) A human snoRNA with microRNA-like functions. *Mol. Cell* 32, 519–528
85. Brameier, M. *et al.* (2011) Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res.* 39, 675–686
86. Macias, S. *et al.* (2012) DGCR8 HITS-CLIP reveals novel functions for the microprocessor. *Nat. Struct. Mol. Biol.* 19, 760–766
87. Wu, H. *et al.* (2017) The diversity of long noncoding RNAs and their generation. *Trends Genet.* 33, 540–552
88. Quinn, J.J. and Chang, H.Y. (2016) Unique features of long non-coding RNA biogenesis and function. *Nat. Rev. Genet.* 17, 47–62
89. Kopp, F. and Mendell, J.T. (2018) Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172, 393–407
90. Lemay, J.F. and Bachand, F. (2015) Fail-safe transcription termination: because one is never enough. *RNA Biol.* 12, 927–932
91. Djebali, S. *et al.* (2012) Landscape of transcription in human cells. *Nature* 489, 101–108
92. Casanal, A. *et al.* (2017) Architecture of eukaryotic mRNA 3′-end processing machinery. *Science* 358, 1056–1059
93. Mischo, H.E. and Proudfoot, N.J. (2013) Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast. *Biochim. Biophys. Acta* 1829, 174–185
94. Shi, Y. and Manley, J.L. (2015) The end of the message: multiple protein-RNA interactions define the mRNA polyadenylation site. *Genes Dev.* 29, 889–897
95. Mason, P.B. and Struhl, K. (2005) Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. *Mol. Cell* 17, 831–840
96. Gromak, N. *et al.* (2006) Pause sites promote transcriptional termination of mammalian RNA polymerase II. *Mol. Cell. Biol.* 26, 3986–3996
97. Schwalb, B. *et al.* (2016) TT-seq maps the human transient transcriptome. *Science* 352, 1225–1228
98. Baejen, C. *et al.* (2017) Genome-wide analysis of RNA polymerase II termination at protein-coding genes. *Mol. Cell* 66, 38–49
99. Baillat, D. *et al.* (2005) Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* 123, 265–276
100. Neuenkirchen, N. *et al.* (2008) Deciphering the assembly pathway of Sm-class U snRNPs. *FEBS Lett.* 582, 1997–2003
101. Siprashvili, Z. *et al.* (2016) The noncoding RNAs SNORD50A and SNORD50B bind K-Ras and are recurrently deleted in human cancer. *Nat. Genet.* 48, 53–58
102. Zhou, F. *et al.* (2017) AML1-ETO requires enhanced C/D box snoRNA/RNP formation to induce self-renewal and leukaemia. *Nat. Cell Biol.* 19, 844–855
103. Kishore, S. and Stamm, S. (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* 311, 230–232
104. Kishore, S. *et al.* (2010) The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. *Hum. Mol. Genet.* 19, 1153–1164
105. Vitali, P. *et al.* (2010) Long nuclear-retained non-coding RNAs and allele-specific higher-order chromatin organization at imprinted snoRNA gene arrays. *J. Cell Sci.* 123, 70–83
106. Vilborg, A. *et al.* (2015) Widespread inducible transcription downstream of human genes. *Mol. Cell* 59, 449–461
107. Rutkowski, A.J. *et al.* (2015) Widespread disruption of host transcription termination in HSV-1 infection. *Nat. Commun.* 6, 7126
108. Bonev, B. and Cavalli, G. (2016) Organization and function of the 3D genome. *Nat. Rev. Genet.* 17, 772